On approximation with spline generated framelets

Rémi Gribonval, Morten Nielsen

ON APPROXIMATION WITH SPLINE GENERATED FRAMELETS

R. GRIBONVAL AND M. NIELSEN

ABSTRACT. We characterize the approximation spaces associated with the best n-term approximation in $L_p(\mathbb{R})$ by elements from a tight wavelet frame associated with a spline scaling function. The approximation spaces are shown to be interpolation spaces between L_p and classical Besov spaces, and the result coincide with the result for nonlinear approximation with an orthonormal wavelet with the same smoothness as the spline scaling function. We also show that, under certain conditions, the Besov smoothness can be measured in terms of the sparsity of expansions in the wavelet frame, just like the nonredundant wavelet case. However the characterization now holds even for wavelet frame systems that do not have the usually required number of vanishing moments, *e.g.* for systems built through the Unitary Extension Principle, which can have no more than one vanishing moment. Using these results, we describe a fast algorithm that takes as input any function and provides a near sparsest expansion of it in the framelet system as well as approximants that reach the optimal rate of nonlinear approximation. Together with the existence of a fast algorithm, the absence of need for vanishing moments may have an important qualitative impact for applications to signal compression, as high vanishing moments usually introduce Gibbs-type phenomenon (or "ringing" artifacts) in the approximants.

1. Introduction

We are interested in characterizing the approximation spaces associated with the best n-term approximation in $L_p := L_p(\mathbb{R})$, $1 < p < \infty$, by elements from a tight wavelet frame. For orthonormal and bi-orthogonal wavelets such spaces are known to be characterized by the sparsity of the wavelet coefficients, and they turn out to be (essentially) Besov spaces [15]. Tight wavelet frames are different from orthonormal wavelets in one important respect; they are (in general) redundant systems but with the same fundamental structure as wavelet systems.

Most constructions of tight wavelet frames based on a multiresolution structure are spline based since it helps make the construction easy and transparent. In the present paper we will keep up this tradition and only consider wavelet frames in L_2 with an underlying multiresolution structure generated by a B-spline. For such systems we characterize the approximation spaces and prove that they are characterized by the sparsity of the framelet representation. Because the framelet system is redundant, the sparsity is defined by searching over all possible representations and choosing the "most sparse" one. As we will show, this leads to a characterization of the approximation spaces (essentially) in terms of Besov spaces just like the nonredundant wavelet case. However, the characterization now holds even for framelet systems that do not have the number of vanishing moments usually required by the theory, because the sparsity is no longer defined in terms of the inner products with the framelets, or analysis coefficients, but in terms of the optimized synthesis coefficients.

Date: January 17, 2002.

Key words and phrases. Wavelet frame, framelets, spline multiresolution, bi-orthogonal spline wavelet, nonlinear approximation, redundant system, adaptive sparse expansions, direct and inverse theorems, Besov spaces.

Research supported by NSF grant no. DMS-9872890.

Characterization of approximation spaces for framelet systems has not only purely theoretical interest, but will also provide some guidance to which class of functions/objects can be compressed efficiently by (nonlinear) framelet expansions. It has already been demonstrated that wavelets do a good job compressing images [12], but are framelets "better" in any quantitative way ? The results in this paper show that no substantial asymptotic gain can be hoped for (that is to say no improvement of the distortion-rate curve at high bitrates), but there may be a substantial gain for approximations with few framelets (*i.e.* compression with a small bitbudget) compared to approximation with few wavelets. This gain may be quantitative (smaller distortion in L_p norm) as well as qualitative: because we can use framelets with only one vanishing moment, we may obtain approximants which yield less Gibbs phenomenon (*i.e.* less "ringing" artifacts) than with wavelets.

The structure of the paper is as follows: Section 2 contains the basic definitions and results needed for the paper. First, we introduce the framelet systems based on the B-spline multiresolution analysis. In Section 2.3 some basic facts about bi-orthogonal wavelets are discussed and we recall the definition of the Chui-Wang semi-orthogonal wavelet that will play an important role later. In the last part of Section 2 we define the approximation spaces associated with best n-term approximation with elements from the framelet system or from oversampled versions thereof. Abstract smoothness spaces associated with the framelet system are also introduced, and we recall the definition of Besov spaces as well as some elements of interpolation theory.

In Section 3 we prove that a nice bi-orthogonal wavelet can be build in such a way that it has a finite expansion in the twice oversampled framelet system. The wavelet will turn out to be either the Chui-Wang semi-orthogonal spline wavelet or, in some cases, a slight modification thereof.

Section 4 contains direct and inverse estimates for approximation with the framelet system. The estimates are used to completely characterize the approximation spaces associated with best n-term approximation in L_p with elements from the twice oversampled framelet system. The direct estimate (Jackson inequality) is based on the expansion of the Chui-Wang type wavelet in the oversampled framelet system obtained in the previous section, and the inverse estimate (Bernstein inequality) is based on Petrushev's results on approximation with free-knot splines. Similar estimates are also obtained for the framelet system itself (i.e. without oversampling) but these results are limited by the number of vanishing moments of the framelet system which may be only one. Section 4 also contains a characterization of certain Besov spaces in terms of sparsity of the (non-unique) framelet coefficients.

Section 5 contains some more results and some open problems. In particular, we show that in some, but not all cases, the direct estimates in Section 4 hold for the framelet system itself, and not only for the associated twice oversampled system. Finally, there is a conclusion in Section 6 where we discuss practical algorithms and applications to signal compression.

2. Framelets and nonlinear approximation

2.1. **Tight wavelet frames.** We will briefly touch upon some of the main ideas in the construction of multiresolution analysis (MRA) based tight wavelet frames, see [10, 25, 24]. For historical notes on this construction, we refer the reader to [10]. Such MRA based tight wavelet frames are called **framelets**. We begin by introducing some basic notation and general assumptions.

Let $\tau = (\tau^0, \tau^1, \ldots, \tau^L)$ be a vector of 2π-periodic measurable functions with τ^0 the mask of a refinable scaling function ϕ of a MRA $\{V_j\}_{j\in\mathbb{Z}}$. We assume that ϕ satisfies $\lim_{\xi\to 0}\widehat{\phi}(\xi) = 1$ and there exist $0 < A \le B < \infty$ such that $A \le \sum_{k\in\mathbb{Z}} |\widehat{\phi}(\cdot - 2\pi k)|^2 \le B$, i.e. ϕ generates a

Riesz basis of the scaling space V_0 of the MRA. We associate the "wavelets" $\Psi = \{\psi^\ell\}_{\ell=1}^L$ to τ by letting $\widehat{\psi^\ell}(2\xi) = \tau^\ell(\xi)\widehat{\phi}(\xi)$. For any finite set F of functions in L_2 we consider the following dictionaries of functions:

- $X_R(F) := \{2^{j/2}\eta(2^j\cdot -k/R)\}_{j,k\in\mathbb{Z}}^{\eta\in F}$, the integer $R \geq 1$ is an "oversampling ratio".
- $X_\infty(F) := \{2^{j/2}\eta(2^j\cdot -b)\}_{j\in\mathbb{Z},b\in\mathbb{R}}^{\eta\in F}$.
- $\overline{X}(F) := \{a^{1/2}\eta(a\cdot -b)\}_{a,b\in\mathbb{R}}^{\eta\in F}$

We will simplify the notations by writing $X(F)$ instead of $X_1(F)$ and, for single generator systems, $X(\psi)$ instead of $X(\{\psi\})$. The following is the fundamental tool to construct framelets:

Theorem 2.1 (The Oblique Extension Principle (OEP) [10])**.** *Suppose there exists a 2π-periodic function Θ that is non-negative, essentially bounded, continuous at the origin with $\Theta(0) = 1$. If for every $\xi \in [-\pi, \pi]$ and $\nu \in \{0, \pi\}$,*

$$\Theta(2\xi)\tau^0(\xi)\overline{\tau^0(\xi+\nu)} + \sum_{\ell=1}^{L} \tau^\ell(\xi)\overline{\tau^\ell(\xi+\nu)} = \begin{cases} \Theta(\xi), & \nu = 0, \\ 0, & \textit{otherwise}, \end{cases} \tag{2.1}$$

then the wavelet system $X(\Psi)$ defined by τ is a tight wavelet frame.

The system $X(\Psi)$ is usually called the **framelet system generated by** Ψ.

Remark 2.2. Theorem 2.1 can be stated in slightly more generality by introducing the notion of a spectrum for the scaling space V_0 and dropping the requirement that ϕ generates a Riesz basis, see [10], but this more general construction will not be used in this paper.

Remark 2.3. For $\Theta \equiv 1$, Theorem 2.1 reduces to the Unitary Extension Principle (UEP) of Ron and Shen [25].

Conversely, if we are given a tight wavelet frame Ψ associated to $(\tau^0, \ldots, \tau^L)$ we can define the **fundamental function** Θ by:

$$\Theta(\xi) := \sum_{j=0}^{\infty}\sum_{\ell=1}^{L} |\tau^\ell(2^j\xi)|^2 \prod_{m=0}^{j-1} |\tau^0(2^m\xi)|^2, \tag{2.2}$$

which satisfies Property (2.1) and is 2π-periodic, see [10] for more details.

We will now specialize our setup to the special case of framelets based on B-spline generated MRAs.

2.2. **B-spline generated framelets.** Fix $r \geq 1$ an integer, and let ϕ_r be the r-th order cardinal B-spline defined recursively by $\phi_1(x) = \chi_{[0,1)}(x)$ and $\phi_{r+1} = \phi_r * \phi_1$. The function ϕ_r is a scaling function for a MRA $V_j := \mathrm{clos}_{L_2}\{\mathrm{span}\{\phi_r(2^j\cdot -k)\}_{k\in\mathbb{Z}}\}$, $V_j \subset V_{j+1}$ with Fourier transform

$$\widehat{\phi_r}(\xi) = \left(\frac{\sin(\xi/2)}{\xi/2}\right)^r e^{-ir\xi} \tag{2.3}$$

and refinement mask $\tau_r^0(\xi) = m_r^0(e^{-i\xi})$ where

$$m_r^0(z) = \left(\frac{1+z}{2}\right)^r. \tag{2.4}$$

Let $\Psi = \{\psi^\ell\}_{\ell=1}^L$ be a finite set of **compactly supported** generators of a tight wavelet frame. Unless otherwise specified, we will always assume that the frame is associated with the **MRA generated by the B-spline** ϕ_r, and that the masks $(\tau^1, \ldots, \tau^L)$ are **trigonometric**

polynomials. We denote by $m^\ell(z) = \sum_{k\in\mathbb{Z}} a_k^\ell z^k, 0 \le l \le L$, the Laurent polynomial such that $\tau^\ell(\xi) = m^\ell(e^{-i\xi})$. For z on the unit circle, $\overline{z} = z^{-1}$ hence we can write $\overline{m^\ell(z)} = \overline{m^\ell}(z^{-1}) := \sum_{k\in\mathbb{Z}} \overline{a_k^\ell} z^{-k}$. Note that $X_R(\Psi)$ and $\overline{X}(\Psi)$ are unchanged when a framelet ψ^ℓ in Ψ is replaced by one of its shifts $\psi^\ell(\cdot - k)$, for some $k \in \mathbb{Z}$. Hence we can always assume that the masks are pure polynomials with $m^\ell(0) \neq 0$, *i.e.* $a_k^\ell = 0, k \le -1$ and $a_0^\ell \neq 0$. We let

$$G(z) := \gcd\{m^\ell(z), 1 \le \ell \le L\} \tag{2.5}$$

this will be used in Section 3. As ϕ_r is piecewise polynomial with $r+1$ knots and r nonzero polynomial pieces of degree $r-1$, each ψ^ℓ is also piecewise polynomial with at most

$$\kappa = (r+1) \cdot \max \deg(\tau^\ell)_{\ell=1}^L \tag{2.6}$$

knots.

2.3. **Bi-orthogonal B-spline wavelets.** Throughout this paper, we only consider bi-orthogonal wavelet systems $(\psi, \widetilde{\psi})$ that are MRA based, i.e. such that there exists an associated bi-orthogonal scaling function pair $(\phi, \widetilde{\phi})$ that generates a pair of MRA's $\{V_j\}$ and $\{\widetilde{V}_j\}$ for which Mallat's algorithm holds:

$$\sum_{k\in\mathbb{Z}} \langle f, \widetilde{\phi}_{j+1,k}\rangle \phi_{j+1,k} = \sum_{k\in\mathbb{Z}} \langle f, \widetilde{\phi}_{j,k}\rangle \phi_{j,k} + \sum_{k\in\mathbb{Z}} \langle f, \widetilde{\psi}_{j,k}\rangle \psi_{j,k}. \tag{2.7}$$

We refer to [8, 9] for more details on such bi-orthogonal systems.

We first recall Chui and Wang's construction [6] of a bi-orthogonal spline wavelet pair $(\psi_r, \widetilde{\psi}_r)$ based on a scaling function pair $(\phi_r, \widetilde{\phi}_r)$, where ϕ_r is the B-spline of order r and ψ_r has compact support. The ψ_r with minimal support (contained in $[0, 2r-1]$) is given by $\widehat{\psi_r}(2\xi) := m_r(e^{-i\xi})\widehat{\phi_r}(\xi)$, with mask

$$m_r(z) := \left(\frac{1-z}{2}\right)^r \times \sum_{k=0}^{2r-2} \phi_{2r}(k+1)(-z)^k. \tag{2.8}$$

The dual functions are given explicitly by

$$\widehat{\widetilde{\phi}}_r(\xi) := \frac{\widehat{\phi_r}(\xi)}{\sum_{k\in\mathbb{Z}} |\widehat{\phi_r}(\xi + 2\pi k)|^2} \quad \text{and} \quad \widehat{\widetilde{\psi}}_r(\xi) := \frac{\widehat{\psi_r}(\xi)}{\sum_{k\in\mathbb{Z}} |\widehat{\psi_r}(\xi + 2\pi k)|^2}. \tag{2.9}$$

From (2.8) and (2.9) it is easy to verify that ψ_r and $\widetilde{\psi}_r$ both have r vanishing moments and are both splines of order r, hence they are in $C^{r-\varepsilon}(\mathbb{R})$ for every $\varepsilon > 0$. Also from (2.9), $\widetilde{\phi}_r$ is a spline of order r, and in this particular case, $V_j = \widetilde{V}_j$. Moreover, $\widetilde{\psi}_r$ and $\widetilde{\phi}_r$ both have exponential decay which can be verified directly from the expressions in (2.9) and, e.g., [9, Proposition 5.4.1].

It is known [5] that these bi-orthogonal wavelets are *semi-orthogonal*, that is to say $V_{j+1} = V_j \oplus W_j$ with

$$W_j := \operatorname{clos}_{L_2}\{\operatorname{span}\{\psi_r(2^j \cdot -k)\}_{k\in\mathbb{Z}}\} = \operatorname{clos}_{L_2}\{\operatorname{span}\{\widetilde{\psi}_r(2^j \cdot -k)\}_{k\in\mathbb{Z}}\}, \tag{2.10}$$

hence Mallat's algorithm (2.7) holds for this bi-orthogonal system.

2.4. **Nonlinear approximation with dictionaries.** We will be interested in characterizing the approximation spaces associated with the framelet system $X(\Psi)$ and certain oversampled versions of it. We will also consider the corresponding spaces for the system $X(\phi)$. First we need some notation to clarify the definition of the relevant approximation spaces. By a dictionary in L_p we mean any collection of non-zero elements of L_p. The approximation spaces for a general dictionary $\mathcal{D}$ in L_p are defined as follows.

Definition 2.4 (Approximation spaces)**.** For $\mathcal{D}$ a dictionary in L_p we let $\Sigma_n(\mathcal{D})$ be the set of all functions S of the form $S = \sum_{I\in\Lambda_n} c_I g_I$, where each $g_I \in \mathcal{D}$ and $\#\Lambda_n \leq n$. The error of the best n-term approximation of $f \in L_p$ from $\Sigma_n(\mathcal{D})$ is

$$\sigma_n(f,\mathcal{D})_p = \inf_{S\in\Sigma_n(\mathcal{D})} \|f - S\|_p,$$

and the approximation space $\mathcal{A}_q^\gamma(L_p,\mathcal{D})$ is defined by

$$|f|_{\mathcal{A}_q^\gamma(L_p,\mathcal{D})} := \left(\sum_{n=1}^\infty \left(n^\gamma \sigma_n(f,\mathcal{D})_p\right)^q \frac{1}{n}\right)^{1/q} < \infty,$$

and (quasi)normed by $\|f\|_{\mathcal{A}_q^\gamma(L_p,\mathcal{D})} = \|f\|_p + |f|_{\mathcal{A}_q^\gamma(L_p)}$, for $0 < q,\gamma < \infty$ with the ℓ_q norm replaced with the sup-norm when $q = \infty$.

One of the nice properties of orthonormal wavelets is that there is an equivalence between the rate of best n-term wavelet approximation of a given function, the sparsity of the wavelet representation of that function, and the smoothness of that function. The same is true for framelet systems of the type considered in this paper, however the relation between rate of approximation, sparsity and smoothness is slightly different. Since the framelet system is redundant, it is not immediately clear what is meant by a sparse representation. As will be shown in section 4 we have to search all possible representations of a given function and choose the one that is the "most sparse", and this representation will correspond to smoothness measured in the Besov (quasi)norm.

Definition 2.5 (Abstract smoothness spaces)**.** For $\mathcal{D} = \{g_k, k \in \mathbb{Z}\}$ a countable **quasi-normalized** dictionary in L_p, $\tau \in (0,\infty)$ and $q \in (0,\infty]$, we let $\mathcal{K}_q^\tau(L_p,\mathcal{D},M)$ denote the set

$$\operatorname{clos}_{L_p}\left\{f \in L_p \mid \exists\Lambda \subset \mathbb{N}, |\Lambda| < \infty,\, f = \sum_{k\in\Lambda} c_k g_k,\; \|\{c_k\}\|_{\ell_q^\tau} \leq M\right\}.$$

Then we define

$$\mathcal{K}_q^\tau(L_p,\mathcal{D}) := \bigcup_{M>0} \mathcal{K}_q^\tau(L_p,\mathcal{D},M),$$

with

$$|f|_{\mathcal{K}_q^\tau(L_p,\mathcal{D})} = \inf\{M : f \in \mathcal{K}_q^\tau(L_p,\mathcal{D},M)\}.$$

Remark 2.6. By abuse of notation we will denote $\mathcal{K}_q^\tau(L_p,\mathcal{D})$ even when $\mathcal{D}$ is not quasi-normalized, simply meaning that one should replace $\mathcal{D}$ with $\{\lambda(g)g, g \in \mathcal{D}\}$, where $c \leq \|\lambda(g)g\|_p \leq C$ for some constants $0 < c \leq C < \infty$. For example, while $X(\Psi) = \{2^{j/2}\psi^\ell(2^j\cdot -k)\}_{j,k\in\mathbb{Z}}$ is obviously quasi-normalized in L_2, we have to replace it with $\{2^{j/p}\psi^\ell(2^j\cdot -k)\}_{j,k\in\mathbb{Z}}$ for the definition of $\mathcal{K}_q^\tau(L_p,X(\Psi))$ when $p \neq 2$.

Besov spaces are essentially the natural spaces related to nonlinear wavelet approximation, and we will use them extensively in this paper. We recall here their definition.

Definition 2.7 (Besov spaces)**.** Let $\alpha > 0$, $l = [\alpha] + 1$ and

$$\omega_l(f,t)_p = \sup_{0<h\le t} \left\| \sum_{k=0}^{l} \binom{l}{k} (-1)^{l-k} f(\cdot - kh) \right\|_p$$

the l-th order L_p-modulus of smoothness. We let

$$|f|_{B_q^\alpha(L_p)} := \left(\int_0^\infty [t^{-\alpha}\omega_l(f,t)_p]^q \frac{dt}{t} \right)^{1/q}, 0 < q < \infty,$$

with the $L_q(dt/t)$ norm replaced with the sup-norm when $q = \infty$. The Besov space $B_q^\alpha(L_p)$ is defined as

$$B_q^\alpha(L_p) := \{f : \|f\|_{B_q^\alpha(L_p)} := \|f\|_p + |f|_{B_q^\alpha(L_p)} < \infty\}.$$

Remark 2.8. The reader can verify by direct computation that on the Sobolev embedding line, $1/\tau = \alpha + 1/p$, the Besov semi-(quasi)norm behaves like $|f|_{B_\tau^\alpha(L_\tau)} = \beta^{+1/p}\|f(\beta\,\cdot)\|_{B_\tau^\alpha(L_\tau)}$ under dilation by $\beta > 0$. This fact will be crucial in the proof of Proposition 4.4.

2.5. **Elements of interpolation theory.** It is well known that the main tool in the characterization of $\mathcal{A}_s^\gamma(L_p,\mathcal{D})$ comes from the link between approximation theory and interpolation theory (see *e.g.* [16, Theorem 9.1, Chapter 7]). The Jackson inequality

$$\sigma_n(f,\mathcal{D})_p \le Cn^{-\alpha}|f|_{B_\tau^\alpha(L_\tau)}, \qquad \forall f \in B_\tau^\alpha(L_\tau), \forall n \ge 1 \tag{2.11}$$

and the Bernstein inequality

$$|S|_{B_\tau^\alpha(L_\tau)} \le C'n^\alpha\|S\|_p, \qquad \forall S \in \Sigma_n(\mathcal{D}) \tag{2.12}$$

(with some constants C and C' independent of f, S and n) imply respectively the continuous embedding $(L_p, B_\tau^\alpha(L_\tau))_{\gamma/\alpha,q} \hookrightarrow \mathcal{A}_q^\gamma(L_p,\mathcal{D})$ and the converse embedding $\mathcal{A}_q^\gamma(L_p,\mathcal{D}) \hookrightarrow (L_p, B_\tau^\alpha(L_\tau))_{\gamma/\alpha,q}$ for all $0 < \gamma < \alpha$ and $q \in (0,\infty]$. The notation $V \hookrightarrow W$ means that the two (quasi)normed spaces V and W satisfy $V \subset W$ and there is a constant $C < \infty$ such that $\|\cdot\|_W \le C\|\cdot\|_V$.

3. Expanding a bi-orthogonal spline wavelet in the framelet system

We are going to show in this section that there exist a nice bi-orthogonal spline wavelet that can be written as a finite linear combination of framelets from the **twice oversampled system** $X_2(\Psi)$. The following is the main result of this section.

Theorem 3.1. *Assume $X(\Psi)$ is a tight framelet system based on the B-spline of order r, with fundamental function $\Theta \ge c$ for some $c > 0$. There exist coefficients $\{p_k^\ell, 1 \le \ell \le L, 0 \le k \le K-1\}$ such that*

$$\psi(x) := \sum_{\ell=1}^{L}\sum_{k=0}^{K-1} p_k^\ell \psi^\ell(x - k/2) \tag{3.1}$$

and

$$\widehat{\widetilde{\psi}}(\xi) := \frac{\widehat{\psi}(\xi)}{\sum_{k\in\mathbb{Z}} |\widehat{\psi}(\xi + 2\pi k)|^2} \tag{3.2}$$

defines a bi-orthogonal spline wavelet system $(\psi,\widetilde{\psi})$ based on the B-spline pair $(\phi_r,\widetilde{\phi}_r)$ of order r, where both ψ and $\widetilde{\psi}$ have r vanishing moments, ψ has compact support and $\widetilde{\psi}$ has exponential decay.

Moreover, if $X(\Psi)$ is built through the UEP, the above results hold with $\psi = \psi_r$ the B-wavelet of Chui and Wang.

Remark 3.2. It follows that the orthogonal complement W_0 of V_0 in V_1 is contained in the span of the half-integer shifts of the framelets. This is noticeable since the integer shifts $\{\psi^\ell(x-k), 1 \le \ell \le L, k \in \mathbb{Z}\}$ of the framelets are not even in W_0 in general: to be in W_0, they would need to be orthogonal to V_0 which is only possible if they have r vanishing moments. However it is known that in UEP-based framelet systems, at least one framelet has no more than one vanishing moment.

To prove Theorem 3.1 we will need a characterization of $G(z)$. Its proof will appear at the end of this section.

Proposition 3.3. *Let $X(\Psi)$ be a tight framelet system based on the B-spline of order r.*

- *If $X(\Psi)$ is built through the UEP,* i.e $\Theta \equiv 1$, *then*

$$G(z) = 1 - z. \tag{3.3}$$

- *If $\Theta(\xi) \ge c > 0$ for all ξ then*

$$G(z) = (1-z)^n \widetilde{G}(z) \tag{3.4}$$

where $1 \le n \le r$ and $\widetilde{G}(z)$ has no zero on the unit circle.

Remark 3.4. Note that in the first case, $G(z) = 1 - z$ gives the well known fact [10] that when framelets are built from a B-spline by the UEP, at least one of the framelets has exactly one vanishing moment.

Remark 3.5. We conjecture that $\widetilde{G}(z) = 1$ holds for any spline framelet system. This can be checked manually on several examples [10, Examples 2.18, 2.19, A.1.a].

Using Proposition 3.3 we are now in a situation to prove Theorem 3.1.

Proof. (Theorem 3.1) We take

$$m(z) := \widetilde{G}(z)\widetilde{G}(-z)m_r(z) \tag{3.5}$$

with $m_r(z)$ the mask of Chui and Wang's B-wavelet (see Equation (2.8). One can easily check from Proposition 3.3 that $m(z) = Q(z)G(z)$ for some polynomial $Q(z)$. Moreover, by Bezout's theorem, there exist polynomials $\{q^\ell(z), 1 \le \ell \le L\}$ such that $\sum_{\ell=1}^L q^\ell(z)m^\ell(z) = G(z)$. Hence

$$m(z) = \sum_{\ell=1}^L Q(z)q^\ell(z)m^\ell(z) = \sum_{\ell=1}^L p^\ell(z)m^\ell(z)$$

with some polynomials $p^\ell(z)$. It follows that if we define ψ with mask $m(z)$ based on the scaling function ϕ_r we have

$$\widehat{\psi}(2\xi) := m(e^{-i\xi})\widehat{\phi}_r(\xi) = \sum_{\ell=1}^L p^\ell(e^{-i\xi})\tau^\ell(\xi)\widehat{\phi}_r(\xi) = \sum_{\ell=1}^L p^\ell(e^{-i\xi})\widehat{\psi^\ell}(2\xi).$$

This gives Eq (3.1). There only remains to check that the pair $(\psi, \widetilde{\psi})$ defined by (3.1) and (3.2) is a bi-orthogonal wavelet pair with the desired moments/smoothness/decay properties.

First, if $X(\Psi)$ is built through the UEP, Proposition 3.3 shows that indeed $m(z) = m_r(z)$, hence $\psi = \psi_r, \widetilde{\psi} = \widetilde{\psi}_r$ and the conclusion is reached. Let us now address the general case case where $X(\Psi)$ is built through the OEP with $\Theta \geq c > 0$. As $\widetilde{G}(z)\widetilde{G}(-z)$ is an even polynomial with no zero on the unit circle, we can write it

$$\widetilde{G}(z)\widetilde{G}(-z) = P(z^2)$$

with $P(z)$ some polynomial with no zero on the unit circle, and we obtain $m(z) = P(z^2)m_r(z)$. It follows that $\widehat{\psi}(2\xi) = P(e^{-2i\xi})m_r(e^{-i\xi})\widehat{\phi}_r(\xi) = P(e^{-2i\xi})\widehat{\psi}_r(2\xi)$ *i.e.*

$$\widehat{\psi}(\xi) = P(e^{-i\xi})\widehat{\psi}_r(\xi).$$

It easily follows that

- ψ is a finite linear combination of integer shifts of ψ_r, hence it is a spline with r vanishing moments and compact support.
- $\psi_r(\cdot) = \sum_n a_n \psi(\cdot - n)$ with $\{a_n\} \in \ell_2$ the Fourier series of $1/P(e^{-i\xi})$, hence

$$\operatorname{clos}_{L_2}\{\operatorname{span}\{\psi(2^j \cdot -k)\}_{k\in\mathbb{Z}}\} = \operatorname{clos}_{L_2}\{\operatorname{span}\{\psi_r(2^j \cdot -k)\}_{k\in\mathbb{Z}}\}$$

  equals W_j, the orthogonal complement of V_j in V_{j+1} (see Eq. (2.10)), and Mallat's algorithm, (2.7), holds for the bi-orthogonal system $(\psi, \widetilde{\psi})$.
- The dual wavelet $\widetilde{\psi}$ satisfies $\widehat{\widetilde{\psi}}(\xi) = \widehat{\widetilde{\psi}_r}(\xi)/\overline{P(e^{-i\xi})}$, from which it is easy to check that $\widetilde{\psi}$ is a spline of order r with r vanishing moments and exponential decay.

□

Remark 3.6. By using the Euclidean algorithm, one can explicitly build polynomials $q^\ell(z)$ which satisfy the Bezout relation with degrees $\deg q^k \leq \max_{1\leq\ell\leq L}\{\deg m^\ell - 1\}$. Let us take the example of the UEP case : as $\deg p^\ell = \deg q^\ell + \deg Q = \deg q^\ell + \deg m_r - 1 = \deg q^\ell + 3(r-1)$, we get an upper estimate for the smallest K in Theorem 3.1

$$K - 1 = \max_\ell \deg p^\ell \leq 3(r-1) + \max_{1\leq\ell\leq L}\{\deg m^\ell - 1\}.$$

Let us now turn to a proof of Proposition 3.3. We will use the following result, which is easily derived from [10, Proposition 1.7] (see also [25]).

Lemma 3.7. *Let $X(\Psi)$* **any** *tight framelet system (not necessarily based on the B-splines), with combined MRA mask τ a trigonometric polynomial. Then there exist a rational function $T(z)$, with $T(1) = 1$, such that for (almost) all ξ, $T(e^{-i\xi}) = \Theta(\xi)$, and*

$$m^0(z)\overline{m^0}(z^{-1})T(z^2) + \sum_{\ell=1}^{L} m^\ell(z)\overline{m^\ell}(z^{-1}) = T(z) \tag{3.6}$$

$$m^0(z)\overline{m^0}(-z^{-1})T(z^2) + \sum_{\ell=1}^{L} m^\ell(z)\overline{m^\ell}(-z^{-1}) = 0. \tag{3.7}$$

Proof of Proposition 3.3. Let $z_0 \neq 0$ be a zero of order n of $G(z)$. From Eq. (3.7) we have

$$m_r^0(z_0)\overline{m_r^0}(-z_0^{-1})T(z_0^2) = 0,$$

which gives three possibilities

(1) $\overline{m_r^0}(-z_0^{-1}) = 0$, *i.e.* $z_0 = 1$. Then $(1-z)^n(1+z^{-1})^n$ is a factor of $G(z)G(-z^{-1})$ and by Eq. (3.7) it is a factor of $m_r^0(z)\overline{m_r^0}(-z^{-1})T(z^2)$. As $T(1) = 1$, this implies $0 \leq n \leq r$.

In the case $\Theta \equiv 1$ we get $n \le 1$ by observing that $(1-z)^n(1-z^{-1})^n$ is a factor of $G(z)G(z^{-1})$ hence, by Eq. (3.6), it is a factor of

$$\begin{aligned} 1 - m_r^0(z)\overline{m_r^0}(z^{-1}) &= 1 - \left(\frac{(1+z)(1+z^{-1})}{4}\right)^r \\ &= \left(1 - \frac{(1+z)(1+z^{-1})}{4}\right)\sum_{k=0}^{r-1}\left(\frac{(1+z)(1+z^{-1})}{4}\right)^k \\ &= \frac{(1-z)(1-z^{-1})}{4}\sum_{k=0}^{r-1}\left(\frac{(1+z)(1+z^{-1})}{4}\right)^k. \end{aligned}$$

(2) $m_r^0(z_0) = 0$, *i.e.* $z_0 = -1$. Then, from Eq. (3.6), $\Theta(\pi) = T(-1) = m_r^0(-1)\overline{m_r^0(-1)}T(1) = 0$, which contradicts the assumption $\Theta \equiv 1$ (resp. $\Theta > 0$). Hence $z_0 = -1$ is not a root of $G(z)$.

(3) $T(z_0^2) = 0$. This is excluded in the case $\Theta \equiv 1$, and implies $|z_0| \ne 1$ in the case $\Theta(\xi) > 0$.

So far we know that $G(z) = (1-z)^n$, where $0 \le n \le 1$ in the case $\Theta \equiv 1$, and $G(z) = (1-z)^n\widetilde{G}(z)$, where $0 \le n \le r$ and $\widetilde{G}(z)$ has no zero on the unit circle, in the case $\Theta > 0$. It is simple to check that $n \ge 1:$ since $\sum_{\ell=1}^{L}|m^\ell(1)|^2 = 1 - T(1)|m_r^0(1)|^2 = 0$ we have $m^\ell(1) = 0$ for $1 \le \ell \le L$, hence $1-z$ is a common divisor of the family $\{m^\ell(z), 1 \le \ell \le L\}$. □

4. Jackson and Bernstein inequalities

In this section we are concerned with Jackson and Bernstein inequalities for the Framelet systems and for certain oversampled versions of the systems. We have already noticed, see Equations (2.11) and (2.12), that such inequalities will allow us to characterize the approximation spaces associated with the framelet systems. Let us now state the main result, the proof will appear later in this section.

Theorem 4.1. *Let $X(\Psi)$ be a framelet system based on the B-spline MRA of order r. We have, with equivalent norms, for $0 < \gamma < \alpha < r$, $1 < p < \infty$, $1/\tau = \alpha + 1/p$, and $0 < q \le \infty$,*

$$\mathcal{A}_q^\gamma(L_p, X_R(\phi_r)) = (L_p, B_\tau^\alpha(L_\tau))_{\gamma/\alpha,q}, \quad 1 \le R \le \infty, \tag{4.1}$$

$$\mathcal{A}_q^\gamma(L_p, \overline{X}(\phi_r)) = (L_p, B_\tau^\alpha(L_\tau))_{\gamma/\alpha,q}. \tag{4.2}$$

Moreover, assuming that $\Theta \ge c > 0$, we have

$$\mathcal{A}_q^\gamma(L_p, X_{2R}(\Psi)) = (L_p, B_\tau^\alpha(L_\tau))_{\gamma/\alpha,q}, \quad 1 \le R \le \infty, \tag{4.3}$$

$$\mathcal{A}_q^\gamma(L_p, \overline{X}(\Psi)) = (L_p, B_\tau^\alpha(L_\tau))_{\gamma/\alpha,q}, \tag{4.4}$$

and

$$B_\tau^\alpha(L_\tau) = \mathcal{K}_\tau^\tau(L_p, X_2(\Psi)). \tag{4.5}$$

For Theorem 4.1, the smoothness parameter α is not limited by the number N of vanishing moments of the generators of the framelet system which may be only $N = 1$ if the system is build with the UEP. If we assume this type of restriction on α we have the following Theorem on the framelet system $X(\Psi)$ instead of the twice oversampled system $X_2(\Psi)$ (*i.e.* Eqs. (4.6) and (4.7) complete Eqs. (4.3) and (4.5)).

Theorem 4.2. *Let $X(\Psi)$ be a framelet system based on the B-spline MRA of order r. Suppose each of the generators of the framelet system has at least $N \geq 1$ vanishing moments, then for $0 < \gamma < \alpha < \min\{r, N\}$, $1 < p < \infty$, and $1/\tau = \alpha + 1/p$,*

$$\mathcal{A}_q^\gamma(L_p, X_{2R+1}(\Psi)) = (L_p, B_\tau^\alpha(L_\tau))_{\gamma/\alpha,q}, \quad 0 \leq R \leq \infty. \tag{4.6}$$

and

$$B_\tau^\alpha(L_\tau) = \mathcal{K}_\tau^\tau(L_p, X(\Psi)). \tag{4.7}$$

We will now indicate how the structure of the proof of Theorems 4.1 and 4.2 will be. First we notice that the following result holds true.

Proposition 4.3. *Let Ψ be any finitely generated framelet system based on the scaling function ϕ. We have the following continuous embeddings, for $0 < \alpha < \infty$, $1 < p < \infty$, $0 < q \leq \infty$, and $1 \leq R \leq \infty$,*

$$\begin{array}{ccc} \mathcal{A}_q^\alpha(L_p, X_R(\Psi)) & \hookrightarrow & \mathcal{A}_q^\alpha(L_p, \overline{X}(\Psi)) \\ \downarrow & & \downarrow \\ \mathcal{A}_q^\alpha(L_p, X_R(\phi)) & \hookrightarrow & \mathcal{A}_q^\alpha(L_p, \overline{X}(\phi)) \end{array}$$

Proof. Follows from the fact that each framelet has a finite expansion $\psi^\ell(x) = \sum_{k\in\mathbb{Z}} a_k^\ell \phi(2x - k)$ in terms of the scaling function ϕ. □

From Proposition 4.3, together with the easy fact that $\mathcal{A}_q^\alpha(L_p, X_2(\Psi)) \hookrightarrow \mathcal{A}_q^\alpha(L_p, X_{2R}(\Psi))$ for $1 \leq R \leq \infty$, we see that it suffices to obtain a Bernstein inequality for the system $\overline{X}(\phi_r)$ and a Jackson inequality for the system $X_2(\Psi)$, to obtain Eqs. (4.2), (4.3), and (4.4), and Eq (4.1) with even R. To obtain (4.1) with odd R and (4.6), respectively, we use the fact that $\mathcal{A}_q^\alpha(L_p, X(F)) \hookrightarrow \mathcal{A}_q^\alpha(L_p, X_{2R+1}(F))$ for $1 \leq R \leq \infty$ and $F \in \{\{\phi_r\}, \Psi\}$. Thus, we only need to prove a Jackson inequality for $X(\phi)$ and $X(\Psi)$, respectively.

In Section 4.1 we will obtain a Bernstein inequality for the system $\overline{X}(\phi)$ (Proposition 4.4). In Section 4.2 we prove a Jackson inequality for $X(\phi)$ (Proposition 4.6) that will give (4.1) for all R and (4.2). In Section 4.3 we prove a Jackson inequality for $X_2(\Psi)$ under the assumption that $\Theta \geq c > 0$ (Proposition 4.10), which gives (4.3) and (4.4). We will prove Eq. (4.5) in Section 4.11 (Proposition 4.13).

Finally, Section 4.15 is devoted to the proof of Eqs. (4.6) and (4.7), which require some vanishing moments on the framelets. First we prove a Jackson inequality for the framelet system (Proposition 4.15) that gives (4.6), then we get (4.7) by a "sandwich" argument.

4.1. **Bernstein inequalities.** We have the following general Bernstein inequality for different dictionaries derived from the framelet system.

Proposition 4.4. *Let $0 < \alpha < r$ and $1 < p < \infty$. There exists a constant $C < \infty$ such that the Bernstein inequality*

$$|S|_{B_\tau^\alpha(L_\tau)} \leq C n^\alpha \|S\|_p, \qquad 1/\tau = \alpha + 1/p,$$

holds for $S \in \Sigma_n(\mathcal{D})$, $n \geq 1$ where $\mathcal{D} \in \{X_R(\Psi), \overline{X}(\Psi), X_R(\phi), \overline{X}(\phi)\}$, $1 \leq R \leq \infty$.

First we prove the following Lemma:

Lemma 4.5. *Fix $\alpha > 0$. For any interval $I = [-A, A]$ there exist $0 < \delta$ and $C < \infty$ such that for $f \in B_p^\alpha(L_p(I))$ with support in $[-\delta, \delta]$, $1 < p < \infty$, $0 < q \leq \infty$,*

$$\|f\|_{B_q^\alpha(L_p(\mathbb{R}))} \leq C \|f\|_{B_q^\alpha(L_p(I))}.$$

Proof. Expand f in a wavelet system on $[-A, A]$ based on a compactly supported wavelet with sufficient smoothness as explained in [7]. If δ is small enough (depends on the support diameter of the wavelet) then f will not meet any of the boundary wavelets and the bounded extension operator $\mathcal{E} : B^\alpha_q(L_p(I)) \to B^\alpha_q(L_p(\mathbb{R}))$ defined in [7] is the identity on f. □

Proof. (Alternative proof for $p = q$): We can assume that $I = [-2, 2]$. Let $\eta \in C^\infty(\mathbb{R})$ be a function with $\mathrm{supp}(\eta) \subset [-2, 2]$, $\eta = 1$ on $[-1/4, 1/4]$ and $\sum_k \eta(x - k) = 1$. By the Fubini property of the space $B^\alpha_p(L_p(\mathbb{R}))$ we have, for any $g \in B^\alpha_p(L_p(\mathbb{R}))$ [27],

$$\|g\|_{B^\alpha_p(L_p(\mathbb{R}))} \asymp \left(\sum_{m\in\mathbb{Z}} \|g\eta(\cdot - m)\|^p_{B^\alpha_p(L_p(\mathbb{R}))} \right)^{1/p}.$$

Let $f \in B^\alpha_p(L_p[-2, 2])$ with $\mathrm{supp}(f) \subset [-1/4, 1/4]$. Thus for $g \in B^\alpha_p(L_p(\mathbb{R}))$ with $g|_{[-2,2]} = f$ it follows that $\|g\|_{B^\alpha_p(L_p(\mathbb{R}))} \geq C\|g\eta(\cdot)\|_{B^\alpha_p(L_p(\mathbb{R}))} = \|f\|_{B^\alpha_p(L_p(\mathbb{R}))}$. The result can then be deduced directly from the definition:

$$\|f\|_{B^\alpha_p(L_p[-2,2])} := \inf_{g\in B^\alpha_p(L_p(\mathbb{R}));\, g|_{[-2,2]}=f} \|g\|_{B^\alpha_p(L_p(\mathbb{R}))}.$$

□

We are ready to prove Proposition 4.4.

Proof. Let δ be as in Lemma 4.5 for $I = [-2, 2]$ and fix $S(x) = \sum_{\lambda\in\Lambda} c_\lambda h_\lambda(x) \in \Sigma_n(\mathcal{D})$. Choose $\beta \geq 1$ such that $g(x) = S(\beta x)$ has support contained in $[-\delta, \delta]$ (the size of β depends on the diameter and center of support of S). Notice that $\|g\|_p = \beta^{-1/p}\|S\|_p$ and also $|g|_{B^\alpha_\tau(L_\tau(\mathbb{R}))} = \beta^{-1/p}|S|_{B^\alpha_\tau(L_\tau(\mathbb{R}))}$ which holds because of the relation $1/\tau = \alpha + 1/p$. Now g is piecewise polynomial with at most $\kappa \cdot n$ knots (κ is defined in Eq. (2.6)) so from the result by P. Petrushev [23] on approximation with free knot splines, we have

$$|g|_{B^\alpha_\tau(L_\tau([-2,2])} \leq C\kappa^\alpha n^\alpha \|g\|_p.$$

Using Lemma 4.5 we write

$$\begin{aligned}
\beta^{-1/p}|S|_{B^\alpha_\tau(L_\tau(\mathbb{R}))} &= |g|_{B^\alpha_\tau(L_\tau(\mathbb{R}))} \\
&\leq \|g\|_{B^\alpha_\tau(L_\tau(\mathbb{R}))} \\
&\leq C\|g\|_{B^\alpha_\tau(L_\tau([-2,2]))} \\
&= C|g|_{B^\alpha_\tau(L_\tau([-2,2]))} + C\|g\|_{L_\tau([-2,2])} \\
&\overset{(*)}{\leq} \tilde{C}\kappa^\alpha n^\alpha \|g\|_p + C'\|g\|_{L_p([-2,2])} \\
&\leq C\kappa^\alpha n^\alpha \|g\|_p \\
&= C\beta^{-1/p}\kappa^\alpha n^\alpha \|S\|_p,
\end{aligned}$$

where in (*) we used $L_p([-2, 2]) \hookrightarrow L_\tau([-2, 2])$. Multiplying by $\beta^{1/p}$ we obtain the desired inequality. □

4.2. **Jackson inequality for the scaling system.** We now turn to the proof of the Jackson inequalities for the different dictionaries derived from the underlying MRA. First we consider a Jackson inequality for the scaling system. Proposition 4.6 below will imply Eqs. (4.1) and (4.2). We sketch a proof of the result using a result by Petrushev [22]. It should be noted that the technique works not only for ϕ_r but for any scaling function satisfying a certain Strang-Fix condition, see [22] for details. An alternative proof of Proposition 4.6 (that works only for ϕ_r) can be found in [15].

Proposition 4.6. *Let $0<\alpha<r$ and $1<p<\infty$. There exists a constant $C<\infty$ such that*
$$\sigma_n(f,X(\phi_r))_p \le Cn^{-\alpha}|f|_{B^\alpha_\tau(L_\tau)}, \qquad \text{for } f\in B^\alpha_\tau(L_\tau),$$
with $1/\tau=\alpha+1/p$.

Proof. By the result of Petrushev [22] there exists a finite set $F\subset\{0,1,\dots,\}\times\mathbb{Z}$ and a corresponding function
$$\theta(x)=\sum_{(j',k')\in F} c_{(j',k')}\phi_r(2^{j'}x-k')$$
for which the system $X(\theta)=\{2^{j/2}\theta(2^jx-k)\}_{j,k\in\mathbb{Z}}$ satisfies the Jackson inequality
$$\sigma_n(f,X(\theta))_p\le Cn^{-\alpha}|f|_{B^\alpha_\tau(L_\tau)}$$
The result follows from
$$\sigma_{\#F\cdot n}(f,X(\phi_r))_p\le\sigma_n(f,X(\theta))_p.$$
□

4.3. **Jackson inequalities for the framelet system.** Now we turn to the Jackson inequalities for the framelet system. First we need a few results about bi-orthogonal B-spline wavelet systems. The results will only be stated in the generality needed for this paper. For more general results we would have to assume vanishing moments of the wavelets, however this is guaranteed here by the fact that $W_0\perp V_0$ and V_0 contains (piecewise) polynomials up to degree r.

Theorem 4.7. *Let $(\psi,\tilde\psi)$ be a bi-orthogonal wavelet system based on the bi-orthogonal B-spline pair $(\phi_r,\tilde\phi_r)$ of Section 2.3. Assume that $|\tilde\psi(x)|\le C(1+|x|)^{-1-\varepsilon}$ for some $\varepsilon>0$, and that ψ has compact support. Then for $0<\alpha<r$, $1<p<\infty$, $0<q\le\infty$,*
$$|f|_{B^\alpha_q(L_p)}\ge C\left(\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}}\left(2^{j(\alpha+1/2-1/p)}|\langle f,\tilde\psi_{j,k}\rangle|\right)^p\right)^{q/p}\right)^{1/q},$$
with the usual modification for $q=\infty$.

Proof (outline). By [15, Eq. (4.27)] we have, for $\alpha<r$, $|f|_{B^\alpha_q(L_p)}\asymp\left(\sum_{j\in\mathbb{Z}}[2^{j\alpha}s_j(f)_p]^q\right)^{1/q}$ with $s_j(f)_p:=\inf_{g\in V_j}\|f-g\|_p$ and $\{V_j\}$ is the spline MRA associated with ϕ_r. Define $P_jf:=\sum_{k\in\mathbb{Z}}\langle f,\tilde\phi_{r,j,k}\rangle\phi_{r,j,k}$ the orthonormal projection onto V_j. Using standard techniques, see e.g. [21, p. 31], it is easy to verify that P_j is bounded on L_p with bound, C_p, independent of j. The fact that P_j is a projection onto V_j ensures that $\|f-P_jf\|_p\le(1+C_p)s_j(f)_p$ from which we get $\|(P_{j+1}-P_j)f\|_p\le 2(1+C_p)s_j(f)_p$. Hence
$$|f|_{B^\alpha_q(L_p)}\ge\tilde C\left(\sum_{j\in\mathbb{Z}}[2^{j\alpha}\|(P_{j+1}-P_j)f\|_p]^q\right)^{1/q}.$$
By Mallat's algorithm, (2.7), we have
$$(P_{j+1}-P_j)f=\sum_{k\in\mathbb{Z}}\langle f,\tilde\psi_{j,k}\rangle\psi_{j,k}.$$
Now, notice that $\psi\in C^{r-\varepsilon}$ for every $\varepsilon>0$ since ψ is a spline of order r. We apply the technique from [21, p. 31] again to get (independent of j)
$$\left\|\sum_{k\in\mathbb{Z}}\langle f,\tilde\psi_{j,k}\rangle\psi_{j,k}\right\|_p\asymp 2^{j(1/2-1/p)}\left(\sum_{k\in\mathbb{Z}}|\langle f,\tilde\psi_{j,k}\rangle|^p\right)^{1/p},$$

from which the claim follows. □

Theorem 4.8. *Let $(\psi, \tilde{\psi})$ be a bi-orthogonal wavelet system based on the bi-orthogonal B-spline pair $(\phi_r, \tilde{\phi}_r)$ of Section 2.3. Assume that $|\tilde{\psi}(x)| \leq C(1+|x|)^{-1-\varepsilon}$ for some $\varepsilon > 0$, and that ψ has compact support. Then for $0 < \alpha < r$, $1 < p < \infty$, $1/\tau = \alpha + 1/p$, there exists a constant $C < \infty$ such that*

$$\sigma_n(f, X(\psi))_p \leq Cn^{-\alpha} |f|_{B^\alpha_\tau(L_\tau)}.$$

Proof. (outline) As above, $\psi \in C^{r-\varepsilon}$ for every $\varepsilon > 0$ since ψ is a spline of order r. Using the technique explained in e.g. [14, §7.4] combined with Theorem 4.7, we notice that the claim of the Theorem will follow if we have $\|f\|_p \asymp \|S(f)\|_p$, $1 < p < \infty$, with $S(f)$ the square function of f given by

$$S(f)(x) = \left(\sum_{j,k \in \mathbb{Z}} |\langle f, \tilde{\psi}_{j,k} \rangle|^2 \chi_{[2^{-j}k, 2^{-j}(k+1)]}(x) \right)^{1/2}.$$

However, the fact that $\|f\|_p \asymp \|S(f)\|_p$, $1 < p < \infty$, under the given hypothesis is well known, see e.g. [13, 28]. □

Remark 4.9. It is possible to generalize Theorems 4.7 and 4.8 to wavelets ψ not based on the B-splines and also to wavelets without compact support by assuming appropriate smoothness and decay of ψ. We refer to, *e.g.*, [15] for more details.

We apply Theorem 4.8 to the bi-orthogonal wavelet systems $(\psi, \widetilde{\psi})$ built using Theorem 3.1. We obtain the following Jackson inequality with $C < \infty$

$$\sigma_n(f, X(\psi))_p \leq Cn^{-\alpha} |f|_{B^\alpha_\tau(L_\tau)},$$

for $\alpha < r$ and $1/\tau = \alpha + 1/p$. Combining with Theorem 3.1 gives the proof of the following result. Notice that the result will imply Eqs. (4.3) and (4.4).

Proposition 4.10. *Assume $X(\Psi)$ is a tight framelet system based on the B-spline of order r, and $\Theta \geq c > 0$. Then the following Jackson inequality holds with $C < \infty$*

$$\sigma_{LKn}(f, X_2(\Psi))_p \leq \sigma_n(f, X(\psi_r))_p \leq Cn^{-\alpha} |f|_{B^\alpha_\tau(L_\tau)} \tag{4.8}$$

for any $f \in L_p$, with $\alpha < r$, $1/\tau = \alpha + 1/p$, and K given by Theorem 3.1.

4.4. **Sparsity of framelet expansions.** We will now prove some more results on the twice oversampled framelet system $X_2(\Psi)$ that will lead to a characterization of the Besov spaces using the smoothness spaces $\mathcal{K}^\tau_q(L_p, X_2(\Psi))$. We begin with the following result: the twice oversampled framelet system $X_2(\Psi)$ is actually a frame in L_2.

Theorem 4.11. *Let $X(\Psi) = \{\psi^\ell\}_{\ell=1}^L$ be* **any** *framelet system (not necessarily based on the B-splines) with compactly supported generators in $C^\beta(\mathbb{R})$ for some $\beta > 0$. Then for $1 \leq p < \infty$ there exists a constant $C_p < \infty$ such that the twice oversampled system $X_2(\Psi)$ satisfies*

$$\left\| \sum_{j,k,\ell,\varepsilon} c^\ell_{j,k,\varepsilon} \frac{\psi^\ell_{j,k,\varepsilon}}{\|\psi^\ell_{j,k,\varepsilon}\|_p} \right\|_p \leq C_p \|\{c^\ell_{j,k,\varepsilon}\}\|_{\ell^p_1},$$

with

$$\psi^\ell_{j,k,\varepsilon}(\cdot) := 2^{j/2} \psi^\ell(2^j \cdot -k - \varepsilon/2)$$

and $\{c^\ell_{j,k,\varepsilon}\}$ any sequence of scalars.

Proof. The case $p = 1$ is trivial. Next we consider the case $p = 2$, and we will show that the system is actually a frame. Let $\{e_\lambda\}_{\lambda\in\Lambda}$ be the canonical basis of $\ell_2(\Lambda)$ with $\Lambda = \mathbb{Z}\times\mathbb{Z}\times\mathbb{Z}_L\times\mathbb{Z}_2$. Define T formally by letting $Te_\lambda = \psi^\ell_{j,k,\varepsilon}(\cdot)$ for $\lambda := (j,k,\ell,\varepsilon)\in\Lambda$. Since $X(\Psi)$ is a frame, we notice that

$$T\left\{\mathrm{clos}_{\ell_2(\Lambda)}\{\mathrm{span}\{e_\lambda|\lambda\in\mathbb{Z}\times\mathbb{Z}\times\mathbb{Z}_L\times\{0\}\}\}\right\} = L_2.$$

Thus, we just have to verify that T is bounded from $\ell_2(\Lambda)$ to $L_2(\mathbb{R})$ (see for example [2]). We fix $\ell\in\mathbb{Z}_L$ and notice that the system

$$X(\psi^\ell(\cdot-1/2)) = \{2^{j/2}\psi^\ell(2^j\cdot -k-1/2)\}_{j,k\in\mathbb{Z}}$$

is a system of **vaguelettes** in the terminology of [20] using that each $\psi^\ell\in C^\beta(\mathbb{R})$. Using [20, Théorème 2, p. 270] we conclude that this subsystem is hilbertian. As this is true for each ℓ we may conclude that T is indeed a bounded mapping of $\ell_2(\Lambda)$ onto L_2.

Take $\{\psi_{j,k}\}$ any orthonormal wavelet system with a compactly supported $C^1(\mathbb{R})$ generator. For each ℓ and $\varepsilon = 0,1$ we consider the integral kernel

$$K^{\ell,\varepsilon}(x,y) = \sum_{j,k\in\mathbb{Z}}\psi^\ell_{j,k,\varepsilon}(x)\overline{\psi_{j,k}(y)}.$$

Standard estimates (see [9, Chap. 9] when $\beta\geq 1$, or [20] for $0<\beta<1$) show that the associated integral operator

$$(T^{\ell,\varepsilon}f)(x) = \int K^{\ell,\varepsilon}(x,y)f(y)\,dy = \sum_{j,k}\langle f,\psi_{j,k}\rangle\psi^\ell_{j,k,\varepsilon}(x)$$

is a Calderon-Zygmund operator (notice that L_2-boundedness follows from the $p=2$ case discussed above). Thus, for each ℓ and ε, $T^{\ell,\varepsilon}$ extends to a bounded operator on L_p, $1<p<\infty$. Given that $\|\psi^\ell\|_p/\|\psi^\ell_{j,k,\varepsilon}\|_p = \|\psi\|_p/\|\psi_{j,k}\|_p$ we obtain

$$\begin{aligned}\left\|\sum_{j,k,\ell,\varepsilon}\frac{c^\ell_{j,k,\varepsilon}}{\|\psi^\ell_{j,k,\varepsilon}\|_p}\psi^\ell_{j,k,\varepsilon}\right\|_p &= \left\|\sum_{\ell,\varepsilon}T^{\ell,\varepsilon}\left(\frac{\|\psi\|_p}{\|\psi^\ell\|_p}\sum_{j,k}\frac{c^\ell_{j,k,\varepsilon}}{\|\psi_{j,k}\|_p}\psi_{j,k}\right)\right\|_p\\ &\leq C_p\sum_{\ell,\varepsilon}\left\|\sum_{j,k}c^\ell_{j,k,\varepsilon}\frac{\psi_{j,k}}{\|\psi_{j,k}\|_p}\right\|_p.\end{aligned}$$

We conclude using the fact that the L_p-normalized wavelet system $\{\psi_{j,k}/\|\psi_{j,k}\|_p\}$ is ℓ^p_1-hilbertian in L_p. □

Remark 4.12. It is known, see [25], that $X_2(\Phi)$ is not a **tight** frame, but a frame nevertheless as we have just demonstrated.

We can now use Lemma 4.11 to obtain the following characterization of the Besov spaces $B^\alpha_\tau(L_\tau)$ in terms of the smoothness spaces $\mathcal{K}^\tau_\tau(L_p,X_2(\Psi))$. Notice that we **do not** require the framelet system to have any prescribed number of **vanishing moments**. This seems to indicate that for a very smooth function f (very smooth compared to the number of vanishing moments of the generators of the framelet system), one should not use the framelet coefficients $\{\langle f,\psi^\ell_{j,k}\rangle\}$ to represent that function but instead optimize the representation $f=\sum c^\ell_{j,k}\psi^\ell_{j,k}$ for sparseness of the coefficients $\{c^\ell_{j,k}\}$. We will further discuss this issue in the Conclusion.

Proposition 4.13. *Let $0<\alpha<r$. For any framelet system $X(\Psi)$ based on the B-spline MRA of order r, assuming that $\Theta\geq c>0$, we have with equivalent norms*

$$B^\alpha_\tau(L_\tau) = \mathcal{K}^\tau_\tau(L_p,X_2(\Psi)),$$

with $1<p<\infty$ and $1/\tau=\alpha+1/p$.

The proof relies on [18, Theorem 3.2] of which we use only the following weak form.

Theorem 4.14 (Gribonval, Nielsen). *For any $1 < p < \infty$, $0 < \tau < p$ and $q \in (0, \infty]$, there is a constant C such that for any normalized dictionary $\mathcal{D}$ in L_p and any $f \in \mathcal{K}_q^\tau(L_p, \mathcal{D})$*

$$\|f\|_{\mathcal{A}_q^\alpha(L_p,\mathcal{D})} \leq C \cdot C_p(\mathcal{D}) \cdot |f|_{\mathcal{K}_q^\tau(L_p,\mathcal{D})}, \qquad 1/\tau = \alpha + 1/p \tag{4.9}$$

with

$$C_p(\mathcal{D}) := \sup_{\|\mathbf{c}\|_{\ell_1^p}=1} \|\sum_k c_k g_k\|_p.$$

Notice that $C_p(\mathcal{D})$ may in general be infinite, but Lemma 4.11 shows that $C_p(\mathcal{D}) \leq C_p < \infty$ for $\mathcal{D} = X_2(\Psi)$ properly normalized in $X = L_p$, $1 \leq p < \infty$.

Proof. From Proposition 4.4 and the above Theorem, we have

$$\mathcal{K}_\tau^\tau(L_p, X_2(\Psi)) \hookrightarrow \mathcal{A}_\tau^\alpha(L_p, X_2(\Psi)) \hookrightarrow B_\tau^\alpha(L_\tau),$$

for $1/\tau = \alpha + 1/p$. For the other embedding, $B_\tau^\alpha(L_\tau) \hookrightarrow \mathcal{K}_\tau^\tau(L_p, X_2(\Psi))$, we take $f \in B_\tau^\alpha(L_\tau)$. Then f has an wavelet expansion $f = \sum_{j,k} c_{j,k}\psi_{j,k}$ where $\psi_{j,k}$ is Chui and Wang's wavelet (see Section 2.3) normalized in L_p, and $\{c_{j,k}\}$ is in ℓ_τ, see Theorem 4.7. But we have seen in Theorem 3.1 that each $\psi_{j,k}$ has a finite expansion in the system $X_2(\Psi)$ with a fixed number of terms that does not depend on j or k. The conclusion follows easily from this fact and this completes the proof of Proposition 4.13. □

4.5. **Where vanishing moments eventually come into play.** We now turn to results where we require some vanishing moments of the generators of the framelet system, unlike the abstract smoothness spaces considered so far. It turns out that in the more restrictive case where the smoothness index α is below the minimum number of vanishing moments of the framelet generators, we can actually characterize the Besov spaces using the framelet coefficients $\{\langle f, \psi_{j,k}^\ell\rangle\}$ instead of the optimized coefficients $\{c_{j,k,\varepsilon}\}$ in the twice oversampled system $X_2(\Psi)$. Thus in such cases smooth functions **compress the framelet coefficients** just as they do for the wavelet coefficients [15]. A consequence of this will be the following Jackson inequality for the framelet system $X(\Psi)$.

Proposition 4.15. *Let Ψ be a framelet system with compactly supported generators each with at least $N \geq 1$ vanishing moments. For $\alpha < \min\{r, N\}$, $1/\tau = \alpha + 1/p$, and $1 < p < \infty$, we have*

$$\sigma_n(f, X(\Psi))_p \leq Cn^{-\alpha}|f|_{B_\tau^\alpha(L_\tau)}, \qquad \textit{for } f \in B_\tau^\alpha(L_\tau).$$

For the proof of Proposition 4.15 we will need Proposition 4.17 below. First we introduce some additional notation. For any framelet system Ψ we define for $0 < p < \infty$, $0 < q \leq \infty$, and $\tau \geq p$,

$$\mathcal{B}_q^\alpha(L_p, X(\Psi)) := \{f \in L_p : \{\langle f, \psi_{j,k}^\ell\rangle\}_{j,k} \in \dot{b}_{p,q}^\alpha,\ \ell = 1, 2, \ldots, L\},$$

where the homogeneous discrete Besov space $\dot{b}_{p,q}^\alpha$ is defined by

$$\left\{\{c_{j,k}\}_{j,k\in\mathbb{Z}} \middle| \|\{c_{j,k}\}\|_{\dot{b}_{p,q}^\alpha} := \left(\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}}\left(2^{j(\alpha+1/2-1/p)}|c_{j,k}|\right)^p\right)^{q/p}\right)^{1/q} < \infty\right\}.$$

We norm $\mathcal{B}_q^\alpha(L_p, X(\Psi))$ by letting

$$\|f\|_{\mathcal{B}_q^\alpha(L_p,X(\Psi))} := \|f\|_p + \sum_{\ell=1}^L \|\{\langle f, \psi_{j,k}^\ell\rangle\}\|_{\dot{b}_{p,q}^\alpha}.$$

Remark 4.16. We notice that in the case where $1/\tau = \alpha + 1/p$, $1 < p < \infty$, $\mathcal{B}_\tau^\alpha(L_\tau, X(\Psi))$ is the collection of functions whose framelet coefficients are contained in ℓ_τ when the coefficient functionals are normalized in $L_{p'}$ with p' the conjugate index to p.

As the notation suggest, the class $\mathcal{B}_q^\alpha(L_p, X(\Psi))$ will coincide with a classical Besov space $B_q^\alpha(L_p)$ for certain values of α. The result is as follows (for the definition of r-regular functions, see [21]).

Proposition 4.17. *Fix $N \geq 1$. Suppose Ψ is a framelet system with each generators being an r-regular function having at least N vanishing moments. Then the following identity holds, with equivalent norms, $0 < p < \infty$, $0 < q \leq \infty$,*

$$\mathcal{B}_q^\alpha(L_p, X(\Psi)) = B_q^\alpha(L_p),$$

provided that $\min\{r, N\} > \max\{J - 1 - \alpha, \alpha\}$ *where* $J = 1/\min\{1, p\}$.

Proof. The embedding $\mathcal{B}_q^\alpha(L_p, X(\Psi)) \hookrightarrow B_q^\alpha(L_p)$ follows immediately from the theory of atomic decomposition of $B_q^\alpha(L_p)$, see [17]. To get the other inclusion, we let $\{\psi_{j,k}^m\}$ be the orthonormal Meyer wavelet on $\mathbb{R}$. Then for $f \in B_q^\alpha(L_p)$ we have an expansion $f = \sum_{j,k} d_{j,k}\psi_{j,k}^m$, with $\{d_{j,k}\} \in \dot{b}_{p,q}^\alpha$, see [14]. We notice that the framelet coefficient $\langle f, \psi_{j',k'}^\ell \rangle$ is given by

$$\langle f, \psi_{j',k'}^\ell \rangle = \sum_{j,k} \langle \psi_{j,k}^m, \psi_{j',k'}^\ell \rangle d_{j,k}. \tag{4.10}$$

For each ℓ we consider the following matrix operator M^ℓ defined on sequences by

$$(M^\ell c)_{j',k'} = \sum_{j,k\in\mathbb{Z}} \langle \psi_{j,k}^m, \psi_{j',k'}^\ell \rangle c_{j,k}.$$

Using the fact that each of the generators of the framelet system is an r-regular function it follows from standard estimates that the matrix M^ℓ is almost diagonal for $\dot{b}_{p,q}^\alpha$ and thus bounded on $\dot{b}_{p,q}^\alpha$ provided $\min\{r, N\} > \max\{J - 1 - \alpha, \alpha\}$, see [17, 20]. It now follows immediately that $f \in \mathcal{B}_q^\alpha(L_p, X(\Psi))$. □

Finally we are in a position to prove the Jackson inequality for the framelet system $X(\Psi)$ (Proposition 4.15).

Proof. (Proposition 4.15) From Lemma 4.11 one can easily check that $C_p(X(\Psi)) < \infty$ with the notation from Theorem 4.14. Applying Theorem 4.14 one gets with $1/\tau = \alpha + 1/p$

$$\sigma_n(f, X(\Psi))_p \leq n^{-\alpha}\|f\|_{\mathcal{A}_\tau^\alpha(L_p, X(\Psi))} \leq Cn^{-\alpha}|f|_{\mathcal{K}_\tau^\tau(L_p, X(\Psi))} \leq Cn^{-\alpha}|f|_{\mathcal{B}_\tau^\tau(L_p, X(\Psi))}$$

It is well know that the B-spline ϕ_r, $r \geq 1$, is r-regular [21, p. 21] so Proposition 4.17 applies to the general setup considered in this paper, and we can conclude using the identification of the framelet smoothness space $\mathcal{B}_\tau^\alpha(L_\tau)$ with a classical Besov space $B_\tau^\alpha(L_\tau)$. □

To conclude this section, the above results together with an easy "sandwich" argument provide the proof of Eq. (4.7) in Theorem 4.2.

5. Additional results and some open questions

In the most important tool we provide (Theorem 3.1) a nice bi-orthogonal wavelet is expanded in the **twice oversampled** framelet system $X_2(\Psi)$. This prevents us from writing results such as Theorem 4.1 directly with $X(\Psi)$, and we need assumptions on the vanishing moments of Ψ to get a complete result in Theorem 4.2. Can we get rid of this oversampling?

5.1. **A direct approach.** The easy way to improve Theorem 3.1 would be to expand the "nice" wavelet ψ of Section 3 in terms of integer shifts of the framelets ψ^ℓ instead of half-shifts. A necessary and sufficient condition for such an expansion to be possible is the existence of polynomials $p^\ell(z)$ such that

$$\widehat{\psi}(\xi) = \sum_{\ell=1}^{L} p^\ell(e^{-i\xi})\widehat{\psi^\ell}(\xi). \tag{5.1}$$

Recall that the high-pass filter $m(z)$ associated with ψ can be written $m(z) = P(z^2)m_r(z)$ for some polynomial P with no zeroes on the unit circle and with $m_r(z)$ the mask of the Chui-Wang wavelet. Thus, (5.1) is equivalent to the existence of polynomials p^ℓ such that

$$P(z^2)m_r(z) = \sum_{\ell=1}^{L} p^\ell(z^2)m^\ell(z). \tag{5.2}$$

We want to study (5.2) in more detail. Using Equation (3.4) we write each mask $m^\ell(z)$ as

$$m^\ell(z) = (1-z)^n(a^\ell(z^2) + zb^\ell(z^2))$$

as well as $m_r(z) = (1-z)^n(A(z^2) + zB(z^2))$. Let $P(z)$ a polynomial with no zero on the unit circle. Now we divide (5.2) by $(1-z)^n$ and split it into its even/odd part to get

$$\begin{aligned} P(z^2)A(z^2) &= \sum_\ell p^\ell(z^2)a^\ell(z^2) \\ P(z^2)zB(z^2) &= \sum_\ell p^\ell(z^2)zb^\ell(z^2) \end{aligned}$$

that is to say

$$P(z)\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = \begin{bmatrix} a^1(z) & \cdots & a^L(z) \\ b^1(z) & \cdots & b^L(z) \end{bmatrix} \begin{bmatrix} p^1(z) \\ \cdots \\ p^L(z) \end{bmatrix}. \tag{5.3}$$

Next we study when (5.3) has a solution in the case $L = 2$.

5.2. **The 2×2 case.** Let us consider the case $L = 2$. Any solution of the 2×2 case will also satisfy

$$P(z)\begin{bmatrix} C(z) \\ D(z) \end{bmatrix} := P(z)\begin{bmatrix} b^2(z) & -a^2(z) \\ -b^1(z) & a^1(z) \end{bmatrix}\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = \Delta(z)\begin{bmatrix} p^1(z) \\ p^2(z) \end{bmatrix} \tag{5.4}$$

with $\Delta(z) := a^1(z)b^2(z) - b^1(z)a^2(z)$. Let us write $\Delta(z) = \tilde{\Delta}(z)\Delta_0(z)$ where $\tilde{\Delta}(z)$ has no zero on the unit circle while all zeros of Δ_0 are on it (or Δ_0 might be identically 1).

Proposition 5.1. *There is a solution $p^1(z), p^2(z), P(z)$ to (5.3) with $P(z) \neq 0$, $|z| = 1$ if and only if $\Delta_0(z)$ divides* $\gcd\{C(z), D(z)\}$.

Proof. **Necessary condition.** Let $p^1(z), p^2(z), P(z)$ be such a solution. By hypothesis $P(z)$ has no zeroes on the unit circle while $\Delta_0(z)$ has all (if any) of its zeroes on the unit circle, so $\gcd(P, \Delta_0) = 1$. However, $\Delta_0 \big| \gcd(P \cdot C, P \cdot D)$ and it follows easily that $\Delta_0 \big| \gcd(C, D)$.
Sufficient condition. Assuming $\Delta_0(z)$ divides $\gcd\{C(z), D(z)\}$, we can take

$$\begin{aligned} p^1(z) &:= C(z)/\Delta_0(z) \\ p^2(z) &:= D(z)/\Delta_0(z) \\ P(z) &:= \tilde{\Delta}(z). \end{aligned}$$

$\square$

Let us now consider two examples of framelet systems (both taken from [10]). The first example shows that (5.3) has a solution in some cases. The second example will show that there are framelet systems for which (5.3) has no solution.

Example 5.2. [10, Example 2.16] Take $r = 2$, $m^0(z) := m_2^0(z)$, $m^1(z) := -(1-z)^2/4$, and $m^2(z) := -\sqrt{2}(1+z)(1-z)/4$. One can easily check that (up to a constant)

$$\begin{bmatrix} a^1(z) & a^2(z) \\ b^1(z) & b^2(z) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

so $\Delta(z) = 2$ and Proposition 5.1 shows that (5.3) has a solution in this case, i.e. that ψ_2 can be written as a linear combination of *integer* shifts of the framelets.

Example 5.3. [10, Example 2.18] Take $r = 2$, $m_2(z) = (1-z)^2(\alpha_1 - \alpha_2 z + \alpha_3 z^2)$ [with α_1, α_2, and α_3 given by (2.8)], $m^1(z) := -(1-z)^2/4$, and $m^2(z) := -\sqrt{\frac{6}{24}}(1-z)^2(z + 4z^2 + z^3)$. It is easy to obtain

$$\begin{bmatrix} a^1(z) & a^2(z) \\ b^1(z) & b^2(z) \end{bmatrix} = \begin{bmatrix} 1 & 4z \\ 0 & 1+z \end{bmatrix},$$

$A(z) = \alpha_1 - \alpha_3 z$, and $B(z) = \alpha_2$. Consequently, $\Delta(z) = 1 + z$ with $\tilde{\Delta}(z) = 1$ and $\Delta_0(z) = 1 + z$. Also, $C(z) = (1+z)(\alpha_1 - \alpha_3 z) - 4\alpha_2 z$ and $D(z) = \alpha_2$ so $\gcd(C, D) = 1$ which implies that Δ_0 does not divide $\gcd(C, D)$ and the expansion of the type (5.3) is *not* possible for this framelet system.

5.3. **The hypothesis $\Theta \geq c > 0$.** Another question relative to our results is the assumption that the fundamental function of the frame system satisfies $\Theta(\xi) \geq c > 0$ for all ξ. While we do not have a proof that this holds for any spline framelet system, we have a series of indications that it is not a very restrictive assumption.

- We have manually checked on several spline framelet systems [10, Examples 2.18, 2.19, A.1.a, A.1.b, A.2, A.3.a, A.3.b, A.4] that, indeed $\Theta \geq 1$.
- It is easy to prove that $\Theta \geq 1$ for all systems built using the systematic construction of spline framelets of high approximation order (see [10, Lemma 3.4]).

A recent result by A. Ron and Z. Shen [26] relates Θ to the dimension function

$$D_\Psi(\xi) := \sum_{\psi \in \Psi} \sum_{k \in 2\pi\mathbb{Z}} \sum_{j=1}^{\infty} |\hat{\psi}(2^j(\xi + k))|^2$$

of the framelet system. Under some assumptions of the scaling function ϕ of a MRA-based framelet system, we have the relation

$$D_\Psi = \Theta[\hat{\phi}, \hat{\phi}], \tag{5.5}$$

where we used the bracket-product notation $[\widehat{f}, \widehat{g}] := \sum_{k \in \mathbb{Z}} \widehat{f}(\xi + 2\pi k)\widehat{g}(\xi + 2\pi k)$. For the spline MRA, the bracket product is easily seen to be continuous and it is bounded and strictly positive by the stability of the shifts of ϕ, see [26]. Hence the problem becomes the strict positivity of the dimension function, however this seems to be an open question.

6. Conclusion and applications

In this paper we have characterized approximation spaces based on tight spline framelet systems and shown they are essentially Besov spaces, just as the orthogonal wavelet case. Contrary to the wavelet case, our characterization holds with no assumption on vanishing moments of the framelets. This is of special importance since framelets constructed using the Unitary Extension Principle can have no more than one vanishing moment.

One of the interesting and challenging aspects of approximation/expansion with a redundant system such as a framelet system is the non-uniqueness of the expansion of a function f in the system. In general redundant dictionaries, this brings some potentially intractable computational issues [11, 19] as well as a potential increase in the quality of approximation of individual functions. Our results show that for framelet systems

- there is a fast expansion/approximation algorithm that reaches near optimal sparsity and the optimal rate of approximation;
- the gain in approximation power (compared to a wavelet basis) may only be marginal.

6.1. **Fast algorithm for near sparsest framelet expansion.** In a redundant dictionary, finding an expansion $f = \sum_k c_k g_k$ that (approximately) minimizes the ℓ_1 norm $\|\{c_k\}\|_{\ell_1}$ is a computationally intensive problem. Combining modern techniques of numerical linear programming and computational harmonic analysis, it is sometimes possible to (approximately) solve this problem with $\mathcal{O}(N^{3.5})$ elementary operations [3]. In a twice oversampled framelet dictionary there is an $\mathcal{O}(N)$ algorithm for that.

For any given framelet system, the Euclidean algorithm can be used to solve for the Bezout relation that yields the expansion coefficients $\{p_n^\ell\}$ in Equation (3.1). Then, a near sparsest expansion of a function f in the framelet system can be obtained as follows. We put into brackets the computational complexity for a finite dimensional signal of size N.

(1) Using Mallat's algorithm, perform a fast expansion $[\mathcal{O}(N)]$

$$f(x) = \sum_{j,m} \langle f, \widetilde{\psi}_{j,m} \rangle \psi_{j,m}(x)$$

with $(\psi, \widetilde{\psi})$ the "nice" bi-orthogonal wavelet system of Theorem 3.1;

(2) Using Equation (3.1), rewrite the above expansion in terms of framelets $[\mathcal{O}(NLK)]$

$$\begin{aligned} f(x) &= \sum_{j,m} \sum_{\ell=1}^{L} \sum_{n=0}^{K-1} \langle f, \widetilde{\psi}_{j,m} \rangle p_n^\ell 2^{j/2} \psi^\ell(2^j x - m - n/2) \\ &= \sum_{j,k,\ell} \underbrace{\sum_{2m+n=k} \langle f, \widetilde{\psi}_{j,m} \rangle p_n^\ell}_{c_{j,k,\ell}} 2^{j/2} \psi^\ell(2^j x - k/2) \end{aligned}$$

This expansion algorithm **adapts to unknown sparsity** of f just as a wavelet expansion does: this is the essence of the proof of the equality $B_\tau^\alpha(L_\tau) = \mathcal{K}_\tau^\tau(X_2(\Psi))$. It follows that the **thresholding algorithm**, which provides m-term approximants $A_m(f)$ by keeping the m largest coefficients from the latter expansion, yields the optimal rate of approximation. That is to say, for all f and α if

$$\sigma_m(f, X_2(\Psi))_{L_p(\mathbb{R})} = \mathcal{O}(m^{-\alpha}), \qquad m \geq 1$$

then

$$\|f - A_m(f)\|_{L_p(\mathbb{R})} = \mathcal{O}(m^{-\alpha}), \qquad m \geq 1.$$

It should be noted, however, that we do not have the stronger result

$$\|f - A_m(f)\|_{L_p(\mathbb{R})} \leq C \sigma_m(f, X_2(\Psi))_{L_p(\mathbb{R})}$$

as can obviously be seen by taking $f = \psi_{j,k}^\ell$.

If the above algorithm is replaced by thresholding performed on the frame decomposition

$$f = \sum_{j,k,\ell} \langle f, \psi_{j,k}^\ell \rangle \psi_{j,k}^\ell,$$

Proposition 4.17 shows that similar good properties hold, however we do not know if the new algorithm can adapt to sparsity/rates of approximation $\alpha > N$, where N is the number of vanishing moments of the system.

6.2. **Application to compression.** It is commonly claimed that m-term expansion from redundant systems can provide lower distortion than from a basis, and this should be useful for, *e.g.*, image compression. With framelet systems, we have just shown that no substantial gain in the rate of nonlinear approximation can be hoped for, compared to approximation from a standard wavelet basis. However this is only a theoretical result, which is far from a "proof" that nothing can be gained from the use of framelets.

First, even though there is an equivalence of the norms $\|\cdot\|_{\mathcal{A}_q^\alpha(L_p, X_R(\Psi))} \asymp \|\cdot\|_{\mathcal{A}_q^\alpha(L_p,\mathcal{B})}$ with $\mathcal{B}$ a wavelet basis, the value of the constants in this equivalence may be significantly different one from another. In particular, $\sigma_m(f, X_2(\Psi))$ may be significantly smaller than $\sigma_m(f,\mathcal{B})$ in general.

Then, the results in this paper only show that no substantial **asymptotic** gain can be hoped for, but there may be a substantial gain for approximations with few framelets compared to approximation with few wavelets. This gain may be quantitative (smaller error $\sigma_m(f, X_2(\Psi)) \ll \sigma_m, (f,\mathcal{B})$ as well as **qualitative**. Indeed, because we can use framelets with only one vanishing moment, we may obtain approximants which yield less "ringing" artifacts (*i.e.* less Gibbs phenomenon) than with wavelets.

Redundancy is a crucial element here: in wavelet approximation, approximants to a function f are built using the inner products $\langle f, \psi_{j,k}\rangle$. Hence, oscillating wavelets (vanishing moments) with $\langle x^n, \psi_{j,k}\rangle, 0 \le n \le r$ are needed to "kill" the low order terms of the Taylor expansion $f(x) \approx a_0 + a_1 x + a_2 x^2 + \ldots$. With a redundant framelet system, we have more freedom in building approximants. Hence, we may take advantage of the variety of shapes (number of vanishing moments) of the framelets to fit the various types of features that are present in a signal/image.

In practical applications of framelet decomposition, it may be useful to combine

- the above described Mallat+rewrite frame decomposition
- standard tight frame coefficients

and some numerical optimization methods, in order to get an expansion that fully takes advantage of the various "feature detectors".

Acknowledgments. The authors would like to thank Amos Ron and Zuowei Shen for interesting discussions of matters related to this paper during our visits to Madison, WI. We should also mention that Narfi Stefansson's promising numerical results using framelets for compression purposes served as an inspiration for much of this work. Finally, we would like to thank Ronald DeVore for his constant encouragement and support.

References

[1] Colin Bennett and Robert Sharpley. *Interpolation of operators.* Academic Press Inc., Boston, MA, 1988.
[2] P. Casazza. Every frame is a sum of three (but not two) orthonormal bases—and other frame representations. *J. Fourier Anal. Appl.*, 4(6):727–732, 1998.
[3] S. Chen and D.L. Donoho. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, January 1999.
[4] Charles K. Chui. *An introduction to wavelets.* Academic Press Inc., Boston, MA, 1992.
[5] Charles K. Chui. On cardinal spline-wavelets. In *Wavelets and their applications*, pages 419–438. Jones and Bartlett, Boston, MA, 1992.
[6] Charles K. Chui and Jian-Zhong Wang. On compactly supported spline wavelets and a duality principle. *Trans. Amer. Math. Soc.*, 330(2):903–915, 1992.

[7] A. Cohen, W. Dahmen, and R. DeVore. Multiscale decompositions on bounded domains. *Trans. Amer. Math. Soc.*, 352(8):3651–3685, 2000.
[8] Albert Cohen. Biorthogonal wavelets. In Charles K. Chui, editor, *Wavelets: A Tutorial in Theory and Applications*, volume 2, pages 123–152. Academic Press, Boston, MA, 1992.
[9] Ingrid Daubechies. *Ten lectures on wavelets.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
[10] Ingrid Daubechies, Bin Han, Amos Ron, and Zuowei Shen. Framelets: MRA-based constructions of wavelet frames. *Preprint*, 2001.
[11] G. Davis, S. Mallat, and M. Avellaneda. Adaptive greedy approximations. *Constr. Approx.*, 13(1):57–98, 1997.
[12] R. De Vore, B. Jawerth, and B.J. Lucier. Image compression through wavelet transform coding. *IEEE Trans. on Inform. Theory*, 38:719–746, March 1992.
[13] R. A. DeVore, S. V. Konyagin, and V. N. Temlyakov. Hyperbolic wavelet approximation. *Constr. Approx.*, 14(1):1–26, 1998.
[14] Ronald A. DeVore. Nonlinear approximation. In *Acta numerica, 1998*, pages 51–150. Cambridge Univ. Press, Cambridge, 1998.
[15] Ronald A. DeVore, Björn Jawerth, and Vasil Popov. Compression of wavelet decompositions. *Amer. J. Math.*, 114(4):737–785, 1992.
[16] Ronald A. DeVore and George G. Lorentz. *Constructive approximation.* Springer-Verlag, Berlin, 1993.
[17] Michael Frazier and Björn Jawerth. Decomposition of Besov spaces. *Indiana Univ. Math. J.*, 34(4):777–799, 1985.
[18] R. Gribonval and M. Nielsen. Nonlinear approximation with redundant dictionaries. *IMI preprint*, 2002.
[19] Lee K. Jones. The computational intractability of training sigmoidal neural networks. *IEEE Trans. Inform. Theory*, 43(1):167–173, 1997.
[20] Yves Meyer. *Ondelettes et opérateurs. II.* Hermann, Paris, 1990. Opérateurs de Calderón-Zygmund. [Calderón-Zygmund operators].
[21] Yves Meyer. *Wavelets and operators.* Cambridge University Press, Cambridge, 1992. Translated from the 1990 French original by D. H. Salinger.
[22] Pencho Petrushev. Bases consisting of rational functions of uniformly bounded degrees or more general functions. *J. Funct. Anal.*, 174(1):18–75, 2000.
[23] Pencho P. Petrushev. Direct and converse theorems for spline and rational approximation and Besov spaces. In *Function spaces and applications (Lund, 1986)*, pages 363–377. Springer, Berlin, 1988.
[24] Amos Ron and Zuowei Shen. Affine systems in $L_2(\mathbf{R}^d)$. II. Dual systems. *J. Fourier Anal. Appl.*, 3(5):617–637, 1997. Dedicated to the memory of Richard J. Duffin.
[25] Amos Ron and Zuowei Shen. Affine systems in $L_2(\mathbf{R}^d)$: the analysis of the analysis operator. *J. Funct. Anal.*, 148(2):408–447, 1997.
[26] Amos Ron and Zuowei Shen. The wavelet dimension function is the trace of a shift-invariant system. *preprint*, 2001.
[27] Hans Triebel. *The structure of functions.* Birkhäuser Verlag, Basel, 2001.
[28] P. Wojtaszczyk. Wavelets as unconditional bases in $L_p(\mathbb{R})$. *J. Fourier Anal. Appl.*, 5(1):73–85, 1999.

IRISA-INRIA, Campus de Beaulieu, 35042 Rennes CEDEX, France
E-mail address: Remi.Gribonval@inria.fr

Department of Mathematics, University of South Carolina, South Carolina 29208, USA
E-mail address: nielsen@math.sc.edu